Warning Concerning Copyright Restrictions

The Copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyright material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction not be "used for any purposes other than private study, scholarship, or research." If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use," that user may be liable for copyright infringement.

CHAPTER SIX

Forgetting

Consolidation Theory
Interference Theory
Decay Versus Interference in Immediate Memory
Experiment: Buildup and Release of PI
Relative Distinctiveness
CogLab Experiment: Von Restorff Effect
Any Evidence for Decay?
Forgetting as Discrimination
Chapter Summary

> *We are told that the famous Athenian Themistocles was endowed with wisdoms and genius on a scale quite surpassing belief; and it is said that a certain learned and highly accomplished person went to him and offered to impart to him the science of mnemonics, which was then being introduced for the first time; and that when Themistocles asked what precise result that science was capable of achieving, the professor asserted that it would enable him to remember everything; and Themistocles replied that he would be doing him a greater kindness if he taught him to forget what he wanted than if he taught him to remember.*
>
> —Marcus Tullius Cicero

Historically, there have been three main views of forgetting: (1) forgetting is due primarily to initial storage problems; (2) forgetting is due primarily to decay of information between storage and retrieval; and (3) forgetting is due primarily to retrieval failures. Reflecting these three views, in this chapter we examine consolidation theory, interference theory, and a theory based on discrimination.

Consolidation Theory

Consolidation theory is not really a theory of forgetting but rather a theory of why information is not stored in the first place. Müller and Pilzecker (1900) argued that learning is not complete at the time that practice or rehearsal ends. Rather, a process known as *perseveration* occurs for a while afterward, and the longer this perseveratory period is, the stronger or more consolidated the memory trace will be. If anything occurs to interrupt this perseveration, then the trace may not have the chance to become consolidated, and recall will not be possible. This idea has substantial biological plausibility. Hebb (1949) proposed that memory involves persistent changes both (1) in the relationship between neurons (or groups of neurons, which he called cell assemblies), either through structural modification or biochemical events; and (2) within a particular neuron that affect communication with its neighbors (see Chapter 8). An obvious suggestion was that perseveration is short-term memory and the consolidated trace is long-term memory.

Any version of consolidation theory makes three key predictions. First, a period of mental inactivity (or neuronal inactivity) is more conducive to perseveration—and hence to more consolidation—than is a period of activity. Memory should be better, then, following a rest period (such as a nap) than following a period of activity (such as engaging in some distractor task). Second, if the perseveration period is interrupted, then the trace cannot be consolidated and the item will not be stored. Memory should be worse, then, following some event that prevents consolidation. Third, if consolidation is prevented, the item should *never* be recalled because the essential storage phase was not completed.

There is ample evidence for the first of these predictions. Ebbinghaus (1885) noted that forgetting is retarded by sleep or, more precisely, that the rate of forgetting is slowed when a period of sleep occurs between study and test. Jenkins and Dallenbach (1924) performed a more systematic investigation in which they had two subjects live in the lab for a period of time. The subjects learned lists of items and then were asked to recall them after 1, 2, 4, or 8 hours of sleep or wakefulness. As Figure 6.1 shows, subjects were able to recall more of the items following a period of sleep than following a period of wakefulness.

According to consolidation theory, the interpretation is simple: When the perseveration process is interrupted by other mental or neuronal activity, the trace cannot be consolidated. The longer the perseveratory process, the more the information will be consolidated and the more likely it is that it will be remembered. This interpretation is consistent with another experiment by Dallenbach, this time using cockroaches as subjects. Minami and Dallenbach (1946) investigated the effects of inactivity when sleep was not involved. They first conditioned roaches to elicit an avoidance response. Some roaches were then allowed to engage in normal activity, whereas others were wrapped in cotton wool and placed in matchboxes. The idea was to prevent the roaches from moving and to minimize stimulation. These restricted roaches showed greater retention of the avoidance response than did the roaches that were allowed to engage in normal activity.

One example often cited as supporting the second prediction of consolidation theory—that if perseveration is interrupted, then the trace cannot be consolidated and the item will not be stored—is *retrograde amnesia*. With retrograde amnesia, people cannot recall the events that occurred before a trauma. According to consolidation theory, the perseveratory activity was interrupted by the trauma, and thus the consolidation process could not be completed. Therefore, the events that occurred right before the trauma were not stored and could not be remembered.

Most experimental studies designed to test consolidation theory have used rats as subjects and induced amnesia by subjecting the rats to electroconvulsive shock (ECS). One side effect of ECS, in both humans and other animals, is retrograde amnesia. A typical study, and one that is widely cited, was performed by Duncan (1949). He first trained rats using an avoidance procedure. A rat was placed in a two-chambered box and received a mild foot shock if it did not move to the other side of the chamber within 10 seconds. Rats can learn this task rather quickly. Duncan varied the interval between completion of

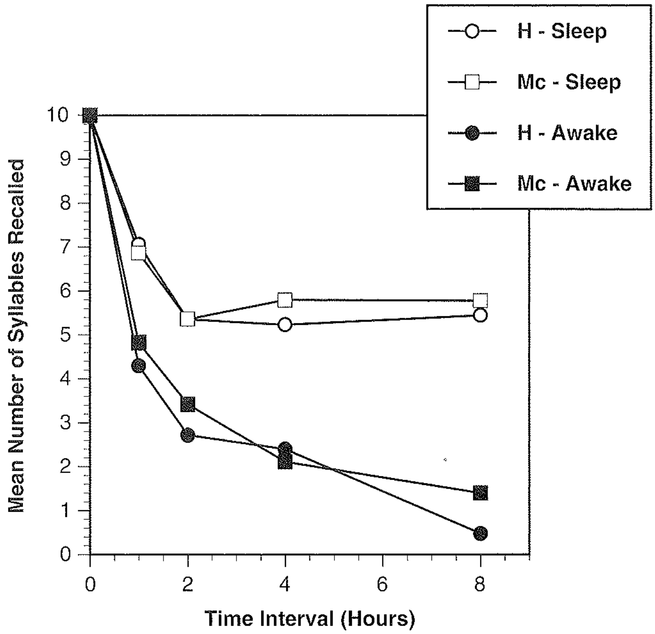

Figure 6.1 Mean number of syllables recalled correctly by two subjects who either slept or remained awake for varying time intervals after learning. SOURCE: From "Obliviscence During Sleep and Waking," by J. G. Jenkins and K. M. Dallenbach, 1924, *American Journal of Psychology, 35,* 605–612. Copyright © 1924 University of Illinois Press. Reprinted with permission.

the avoidance response and the administration of ECS, using intervals from 20 s to 4 hr; a control group received no ECS. Consistent with consolidation theory, the longer the delay between the response and the administration of ECS—or, in terms of the theory, the longer perseveration was allowed to continue—the better was the memory for the learning episode (see Figure 6.2). The estimate of time needed for consolidation was approximately 1 hour.

One problem with this experiment is that it requires an active response on the part of the rat to demonstrate learning. There are two reasons the rat might not move after receiving ECS. First, the animal may have forgotten the avoidance response that it has learned and so does not remember to move to the other side of the chamber to avoid a foot shock. A second possibility however, is that the rats may have inadvertently been subjected to a conditioned fear experiment and may be opting for the lesser of two evils: A move to the other side of the chamber could serve as a cue for ECS, and the animal may prefer the mild foot shock to the severe ECS. Thus, the rat might remain stationary because of amnesia or because of fear.

An experiment by Chorover and Schiller (1965) tested this idea. Rather than require an active response to signify remembering, their experiment required the animal to remain stationary. The experiment involved placing a rat on a raised platform. Because the rat's natural inclination is to avoid exposed, high places, the animal quickly stepped off

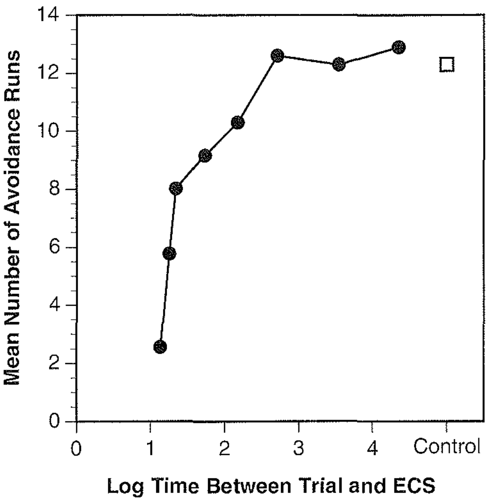

Figure 6.2 Mean number of avoidance runs made by different groups of rats that differed only in the duration of the interval between the learning trial and the ECS. The *x*-axis is in log units. The corresponding values are 20 s, 40 s, 60 s, 4 min, 15 min, 1 hr, 4 hr, and 14 hr. The white square shows performance of animals in a control condition with no ECS. SOURCE: Duncan (1949).

the platform. When this happened, the rat received a mild foot shock; if the rat stayed on the platform, nothing happened. Chorover and Schiller again varied the interval between the learning episode (stepping down) and the administration of ECS. Note that this time, however, if the rat has learned fear of the ECS, it should remain on the platform; if it does not remember the foot shock, it should step down. Whereas Duncan estimated that consolidation takes 1 hour, Chorover and Schiller estimated that consolidation is complete within 10 seconds. If ECS is administered within 10 s of the learning, then there is retrograde amnesia. If ECS is delayed for more than 10 s, then memory for the learning episode will remain. This estimate of 10 s, although obtained on rats, coincided nicely with conceptions of the duration of information in short-term memory.

The third prediction of consolidation theory is that if consolidation is prevented, the item should *never* be recalled. Quartermain, McEwen, and Azmitia (1972) tested this prediction using a procedure similar to that used by Chorover and Schiller. The main difference was that whereas Chorover and Schiller tested all animals 24 hr after the administration of ECS, Quartermain and associates tested some animals 48 and 72 hr after ECS. There was evidence of amnesia in those animals tested 24 or 48 hr after ECS, but no evidence of amnesia after 72 hr. The suggestion is that in the 24- and 48-hr groups, the memory was actually stored but was rendered temporarily unavailable by the ECS; after 72 hr, the memory again became available. This finding directly contradicts the third prediction of consolidation theory: Interrupting the perseveration phase should prevent the storage of information. Quartermain et al. demonstrated that the information is stored; it is just not retrieved.

A final study reinforces this interpretation. Miller and Springer (1972) argued that if the memory is stored but temporarily unavailable, providing a reminder to the animals should remove all traces of amnesia. They followed a procedure similar to that used in the previous studies, but they gave some rats a second foot shock after ECS had been administered. The reminder foot shock was given outside the experimental chamber, so it did not constitute another learning trial. The control group, which did not receive a reminder, showed amnesia; however, the experimental group did not.

Although there is much evidence against a strong form of consolidation theory (Crowder, 1982b; Keppel, 1984; Spear & Riccio, 1994), a weaker form is undoubtedly correct: Some biological changes may indeed affect an organism's ability to store information. This weaker form, though, has yet to provide a link between the physiological processes and the cognitive consequences. As knowledge of psychobiological processes continues to grow, it may again become possible to entertain a serious version of consolidation theory. Until that time, however, several other major theories of forgetting make more accurate predictions and offer more insights into why forgetting occurs.

Interference Theory

A competing view of forgetting was also being developed during the same time as consolidation theory. Rather than suggesting that the information never entered the memory system, it proposed that other information somehow caused interference.

John McGeoch played a central role in the first stages of the development of interference theory. In 1932, he argued that time-based decay theories are not valid scientific explanations of behavior. Specifically, he argued that just as iron rusts over time, memories are forgotten over time, but in neither case is time the causal agent. For iron, it is the process of oxidation that causes rust, and this process must, of necessity, unfold over time. The same is true for forgetting: It must, of necessity, unfold over time, but something else, analogous to oxidation, occurs as time passes. McGeoch suggested that the processing of other material causes interference. Notice that the experiments by Dallenbach reported above can be interpreted as supporting an interference explanation.

Interference can be either retroactive or proactive. To illustrate, consider a hypothetical experiment. In a *paired-associate experiment*, subjects are asked to learn arbitrary pairings of two items, such as *dog* and *roof*, *car* and *desk*, and so forth. at test, they will be prompted with *dog* and asked to respond with the correct answer, *roof*. These item pairs are often represented as A-B, with A standing for the first item, or cue, and B standing for the second item, or target. If subjects learn a second list that contains entirely new items, C-D (such as *phone-coffee*), there will be relatively little interference. In other words, memory for the A-B relationship will not be affected very much by learning the C-D relationship. If, however, the subjects learn a list that reuses the A items but pairs them with new, D items, there will be interference. If *dog-roof* is the A-B pairing, *dog-coffee* might be the A-D pairing. When tested later for the A-B items, the A-D associations will produce what is called *retroactive interference (RI)*, because A-D interferes with something learned earlier. This is illustrated in the top part of Table 6.1. The bottom part of Table 6.1 illustrates the procedure for inducing *proactive interference (PI)*. Note that we are again testing memory of the A-B items, but subjects in the experimental group learn the A-D

Table 6.1 A schematic illustration of two experimental designs in which subjects are tested for the memory of A-B relationships

	Learn First	*Learn Second*	*Tested On*	*Interference*
Control	A-B	C-D	A-B	
Experimental	A-B	A-D	A-B	RI
Control	C-D	A-B	A-B	
Experimental	A-D	A-B	A-B	PI

NOTE: The top design produces retroactive interference; the bottom design produces proactive interference.

associations first. This will cause proactive interference because already-learned information will interfere with newer information.

Although the stimuli and methods described may seem artificial, the principles apply equally as well in real-world settings (see, for example, Bower, 1978). For example, Table 6.2 gives one example of when RI and PI might occur if a person is trying to learn both Italian and Spanish vocabulary. To take another example, as we shall see in Chapter 12, the typical experiment on the effects of misleading information on eyewitnesses is nothing more than a fancy A-B–A-D paradigm (Crowder, 1976): The original material is B, and the misleading information is D.

Underwood (1957) dramatically documented the effects of proactive interference. He combined the results of 14 studies and simply plotted them all on the same graph, shown in Figure 6.3. What he found was that the more previous trials a subject had with a particular task, the lower the performance. That is, if a person has had three previous trials and everything else is held constant, performance on the fourth trial will be worse because of interference from the first three trials. Although some suggestions were made on how to incorporate PI into interference theory, a mechanism was never really specified. This lack of a specific mechanism for PI was one major reason for the diminishing popularity of interference theory in general (Zechmeister & Nyberg, 1982). Despite this omission, however, interference theory has greatly influenced all subsequent theories of forgetting.

Consolidation theory can be seen as one form of interference theory, in which interference occurs during storage. In McGeoch's (1932, 1942) *three-factor theory of forgetting*—the initial main formulation of interference theory—the locus of interference was placed

Table 6.2 A schematic illustration of two experimental designs in which subjects are tested for Spanish vocabulary

	Learn First	*Learn Second*	*Tested On*	*Interference*
Control	Spanish vocabulary	Math	Spanish vocabulary	
Experimental	Spanish vocabulary	Italian vocabulary	Spanish vocabulary	RI
Control	Math	Spanish vocabulary	Spanish vocabulary	
Experimental	Italian vocabulary	Spanish vocabulary	Spanish vocabulary	PI

NOTE: In the top design, learning Italian vocabulary after the Spanish will produce retroactive interference, whereas in the bottom design, learning Italian vocabulary before the Spanish will produce proactive interference.

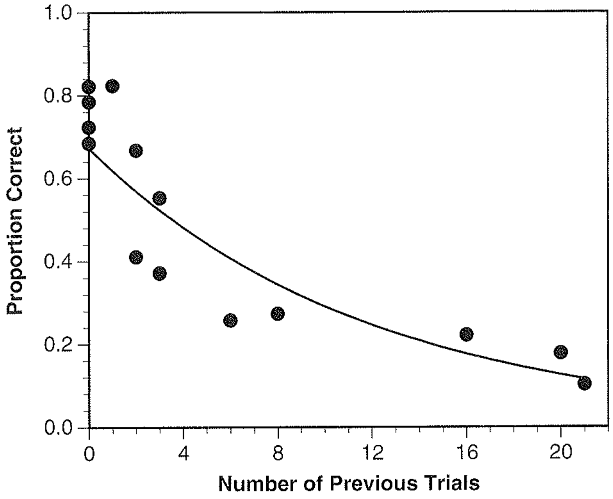

Figure 6.3 Proportion correct plotted as a function of the number of previous trials. Data were collected from 14 different studies.
Source: Underwood (1957).

at retrieval. McGeoch assumed that retrieval failures occur because unwanted memories are retrieved instead of the desired memory. The A-D learning does not erase or remove the memories of the A-B learning, and the memories of the A-B pairing do not decay or fade over time; rather, both B and D coexist as memories for the correct response to A. In McGeoch's theory, there is no forgetting in the sense of loss of information; rather, forgetting is a temporary performance problem. There is no loss of information, only a lapse in retrieving it. McGeoch identified three mechanisms that can cause interference: response competition, altered stimulus conditions, and set.

Response competition occurs when two or more items are potential responses to a memory query. For example, when cued with A-?, both B and D compete as potential responses, and this competition reduces the probability that either one will be offered as a response. Much research was conducted on this aspect of interference theory; Crowder (1976) provides an excellent discussion and review. The main question concerned how best to test for the relative strength of the B and D responses so that different explanations of response competition could be further explicated. For example, Melton and Irwin (1940) used an A-B–A-D paired associate task, much like the one illustrated in Table 6.1. Subjects in the control condition learned to associate a series of items with cues, the A-B learning, and then "rested" for 30 min. Other groups had either 5, 10, 20, or 40 trials learning A-D paired associates, and then all groups were tested for the original A-B material.

If response competition were causing all of the forgetting, then all of the errors in the groups with interpolated learning should be D responses, rather than the desired B responses. Melton and Irwin demonstrated that this was not the case—that the difference in recall between the control and the experimental groups could not be attributed solely to intrusion or competition from D items. They suggested that subjects actually unlearned the original A-B material during the interpolated learning. However, the unlearning hypothesis quickly ran into the same problem that consolidation theory did: It is easy to find situations in which information that is supposed to be completely lost from the memory system reappears (Capaldi & Neath, 1995). For example, Postman and his colleagues (Postman, 1961; Postman, Stark, & Fraser, 1968) demonstrated that specific A-B–A-D associations are not actually lost, as the unlearning hypothesis predicted, but are available when different forms of testing are used.

The second of McGeoch's original three factors was *altered context* or *altered stimulus conditions*. Animal learning theorists distinguish between the nominal stimulus and the functional stimulus. For example, if a tone consistently precedes a mild foot shock, then rats will learn to avoid the shock by moving. From the experimenter's perspective, the tone is the important stimulus and nothing else. From the animal's perspective, however, the tone is only a *nominal* stimulus; additional sources of information also serve as cues. The total of all the cues that the animal attends to is the *functional* stimulus. For example, if the conditioning takes place in a white box, then the functional stimulus may be both a tone and a white environment. When the animal is placed in a black box and the tone is played, the animal may not respond. This does not mean that the animal has forgotten; rather, it could mean that the functional stimulus (white + tone) is no longer present to elicit the response.

Similar effects have long been known with humans as well (Carr, 1925; McGeoch, 1932). As discussed in the previous chapter, Abernathy (1940) demonstrated that students who were tested in the same classroom they were taught in did better than students who were tested in a different classroom. Subsequent research has consistently supported the idea that altered stimulus conditions—or more generally, contextual factors—play a major role in most kinds of forgetting.

The third factor in McGeoch's three-factor theory was set. *Set* can be thought of as a special version of context effects, with the subject using an inappropriate mind-set. The typical example is failing to recognize a friend from college when you happen to run into the person during a visit home. The idea is that, when at home, you are most likely to search through your list of "home friends" to identify a familiar face; because this person is not a "home friend," you find no match. If you had used the correct mind-set—"friends from college"—then you would have recognized the person.

A related phenomenon is known as *part-set cuing*. When people are asked to recall items from a list they have just studied or to produce examples from a well-known category, having a subset (part of a set) of those items available at recall usually does not help and can even impair performance. For example, Slamecka (1968) had subjects hear a list of words twice. The control subjects were given a blank sheet of paper and were asked to write down as many of the words as they could. The experimental subjects were given a sheet of paper that already contained 5 (or 15, or 25) of the words and were asked to supply the remaining 25 (or 15 or 5) items. The data of interest are the proportion of words recalled that were not cues. Slamecka found that the groups that had received part of the set as a cue recalled fewer items than did the control group. That is, the experimental group that received 15 words as cues could produce (on average) only 38% of the remaining words, compared to 50% for the control group. Although there is still no complete account for all of the part-set cuing data (see Nickerson, 1984), one interpretation is that the cues disrupt the normal retrieval strategy.

McGeoch's influence was so profound that no theories of forgetting from long-term memory propose decay as the main explanation. Within the realm of immediate memory, however, the most common view of forgetting has long been decay.

Decay Versus Interference in Immediate Memory

The Brown-Peterson paradigm is one area in which both a decay explanation and one based on interference were offered. Recall that in the Brown-Peterson paradigm, subjects are given three letters and are then asked to count backward (or perform some other distracting activity) for a specified amount of time. The typical finding is that after about 20 seconds of counting backward, very few people can recall the letters. Indeed, many textbooks still cite the results of Peterson and Peterson (1959) as supporting the rapid decay of information in short-term memory. We described the basic results from the Brown-Peterson paradigm in Chapter 3 and saw how they were used as evidence for a separate short-term store and for advocating a major role for decay to explain the loss of information. An enormous amount of research has been conducted in this paradigm precisely because of its importance for theories of memory. This research has generally not favored the decay explanation.

The first major challenge to the decay explanation was reported by Keppel and Underwood (1962). Remember that Underwood (1957) found that, over a large number of studies, subjects' performance decreased as they received more and more trials. Proactive interference could operate in the Brown-Peterson paradigm because all of the consonant trigrams on all of the trials were very similar: The more previous trials that subjects had seen, the worse their performance was likely to be. In fact, Peterson and Peterson were aware of this decline, and analyzed their data to see whether PI was operating. They divided the 48 trials into blocks of 12 and combined performance in the two shortest and the two longest conditions. The proportion of items recalled correctly is displayed in Table 6.3. Peterson and Peterson (1959) concluded that there was no PI because performance actually got better over blocks in the 3 s and 6 s conditions. If there were PI, performance would be expected to get worse.

The only difference between Peterson and Peterson's (1959) analysis and that of Keppel and Underwood (1962) was that the latter study examined the first four trials individually, rather than averaging over the first 12 trials. When they did so, Keppel and Underwood found substantial evidence for proactive interference. On Trial 1, there was no forgetting regardless of how long the subjects counted backward (see Figure 6.4). Forgetting was observed only on subsequent trials, and by as few as 3 trials, performance was quite poor.

The important point is that recall does not differ on the very first trial when tested after 3 s or after 18 s. Equivalent performance has been repeatedly demonstrated when the first trial is analyzed (Baddeley & Scott, 1971; Cofer & Davidson, 1968; Fuchs & Melton, 1974; Gorfein, 1987; Turvey, Brick, & Osborn, 1970; Wright, 1967). A decay explanation, such as the one usually invoked to explain the Peterson and Peterson (1959) results,

Table 6.3 Proportion of items recalled correctly, displayed as blocks of 12 trials for two different sets of delay

	Block 1	*Block 2*	*Block 3*	*Block 4*
3 s and 6 s	0.57	0.66	0.70	0.74
15 s and 18 s	0.08	0.15	0.09	0.12

SOURCE: Data from Peterson & Peterson (1959).

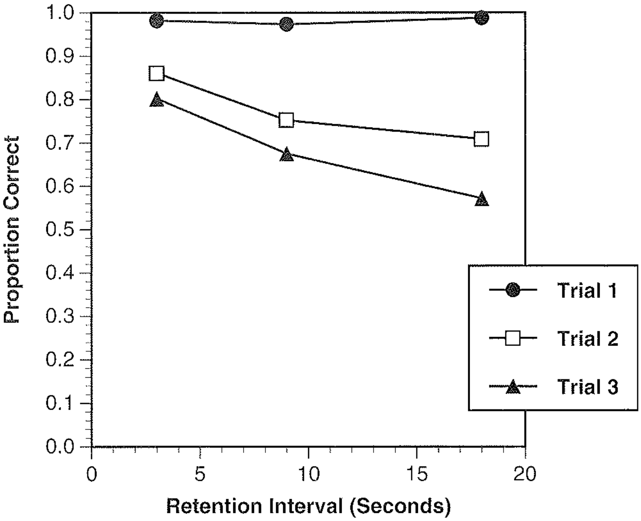

Figure 6.4 Proportion of items recalled correctly in a Brown-Peterson procedure at three different retention intervals on the first, second, and third trials. SOURCE: From "Proactive Inhibition in Short-Term Retention of Single Items," by G. Keppel and B. J. Underwood, 1962, *Journal of Verbal Learning and Verbal Behavior, 1*, 153–161. Copyright © 1962 Academic Press, Inc. Reprinted with permission.

says that the more time that elapses, the more decay will occur. Showing equivalent performance for the 3 s and 18 s groups is inconsistent with this prediction. The reason that Peterson and Peterson (1959) did not observe PI was that it had already occurred within the 12 trials that composed Block 1.

The decay view and the interference view make different predictions. Wickens, Born, and Allen (1963; see also Wickens, 1970, 1972) tested a strong prediction of the PI interpretation of the Brown-Peterson results. Proactive interference will build up when the list items are all highly similar. If the stimulus type is changed, however, there should be less PI. Thus, on the first three trials, the researchers had subjects hear three consonants, then perform some distracting activity, and then recall the consonants. On the fourth trial, half of the subjects were switched from consonants to numbers. Performance in the switched group was much better than in the control group, regardless of the direction of the change (digits to letters or letters to digits). The PI account predicts this increase in performance when the type of stimuli is changed. The decay view, in contrast, cannot predict an increase in performance.

The amount of reduction, or *release from PI*, can be calculated easily (see Figure 6.5). If x is the difference in performance between the experimental and control groups on Trial 4, and y is the difference between the control group on Trial 4 and the control group on Trial 1, then the percent release from PI is equal to $(x/y) \times 100$. If the release is 100%, then performance on the fourth trial is the same as performance on the first trial. If it is close to 0, there is no release.

Most release from PI phenomena are symmetric; that is, the amount of release is the same when switching from consonants to numbers as it is when switching from numbers

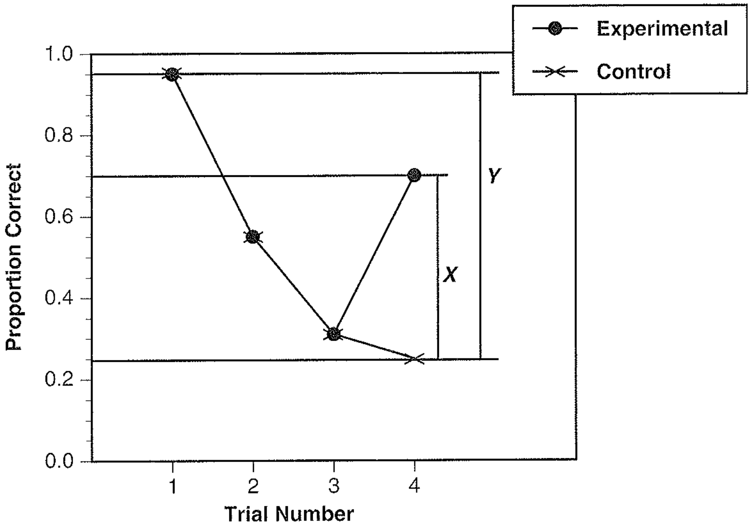

Figure 6.5 Idealized data showing release from proactive interference in the Brown-Peterson paradigm. X is the difference in performance between the experimental group (who have a different type of stimulus material on Trial 4 than on the previous trials) and the control group (who have the same type of stimulus material). Y is the difference in performance between Trial 1 and Trial 4 for the control group.

to consonants. This symmetry indicates it is the change in materials that is crucial, rather than the fact that one type of material is easier to recall than another. Switching from words to numbers gives an 85% release; switching from auditory to visual also gives about an 85% release. Changing taxonomic category (such as from types of cars to Russian novelists) gives a 65% release. Changing the number of syllables gives only a 15% release. Changing from one part of speech to another (verbs to nouns) gives less than a 5% release. One particularly interesting release concerns what is known as the *semantic differential*. If all the words on the first three trials have similar meanings, such as something pleasant (*sunny*, *enjoy*, *happy*), and the words on the last trial mean something different (*kill*, *death*, *illness*), there is about a 55% release from PI. Wickens (1972) reports many such measures.

Release from PI is not limited to the typical laboratory paradigm. Gunter, Berry, and Clifford (1981) demonstrated release from PI when testing subjects' memory for news items taped off the evening news. The news stories were classified as being either domestic or international. Subjects saw several stories from one category, and then either another from the same category or one from the other category. Subjects showed a 70% release from PI. Blumenthal and Robbins (1977) have shown buildup and release of PI when answering questions about prose passages, suggesting that this might affect performance on standardized tests. Dempster (1985) tested this idea directly, demonstrating the buildup of PI on comprehension questions like those used on the ACT and SAT.

Turvey and Egan (1970) changed two dimensions at the same time and found that performance on Trial 4 was equal to performance on Trial 1, regardless of whether the retention interval was 5 s or 15 s. Other results consistent with the proactive interference

Experiment *Buildup and Release of PI*

Purpose: To demonstrate the buildup and release of proactive interference in the Brown-Peterson paradigm.

Subjects: Forty subjects are recommended; ten should be assigned to each group. Two groups receive the same type of items on all four trials, and two groups change to a different type of item on Trial 4.

Materials: Table F in the Appendix lists 32 consonant trigrams. Number trigrams can be obtained from Table B by taking the first three digits, then the second three digits, and so on. Table G lists several names of categories. For each subject, present novel items. There should also be two answer sheets: one for the category items and one for the trigrams. You will also need a stopwatch.

Design: One group of subjects will receive only letters, and one will receive only numbers. For the remaining subjects, one group will receive letters for the first three trials and numbers on the fourth, and the other group will receive numbers for the first three trials and letters on the fourth. Because each subject experiences only one condition, this is a between-subjects design.

Procedure: Inform the subjects that this is an experiment on information processing. Describe the two tasks, and explain how subjects should respond on the answer sheets. Read each trigram aloud, immediately name a category, and ask subjects to name as many examples of that category as they can. Record the examples of the category that the subject produces. After 10 s, ask the subject to report the trigram; then begin the next trial. On the fourth trial, change the type of material for half of the subjects.

Instructions: "This is an experiment on information processing. I will read three items out loud, which I will ask you to recall in order. Before you recall them, I will name a category, and I would like you to name as many examples as you can as quickly as you can. For example, if I said 'dog,' you might say 'poodle, corgi, greyhound, Pekinese, Great Dane, English setter' as quickly as you can. When I signal, you will stop and report the three items in order. Then, I'll give you the next three items and a new category. Any questions?"

Scoring and Analysis: For each subject, count the number of trigrams recalled in order, and compute an average for each group. The results from the two groups that receive only letters or only numbers can be averaged together, and the results from the two groups that experienced a change on the fourth trial can be averaged together. The results should be similar to those plotted in Figure 6.5.

Optional Enhancements: Include more subjects in each group.

SOURCE: Based on an experiment by Wickens (1972).

explanation were reported by Loess and Waugh (1967). If recall is being reduced by PI, then one way to decrease the amount of PI is to separate the trials. Loess and Waugh found that with very long intervals between trials (such as 300 s), there was no difference in performance as a function of trial number, even when the subjects were counting backward for 18 seconds. Subjects were about 60% accurate on Trial 1 and were still around 60% correct on Trial 4. Both results are inconsistent with an explanation that attributes forgetting to the rapid loss of information from short-term memory.

Several studies have tried to pin down the locus of the release from PI effect. An early suggestion was that when a change in stimulus properties occurs, subjects devote extra attention to the novel item. This account has two problems. One is that subjects are often unaware of the changes, particularly with the semantic differential, yet they still

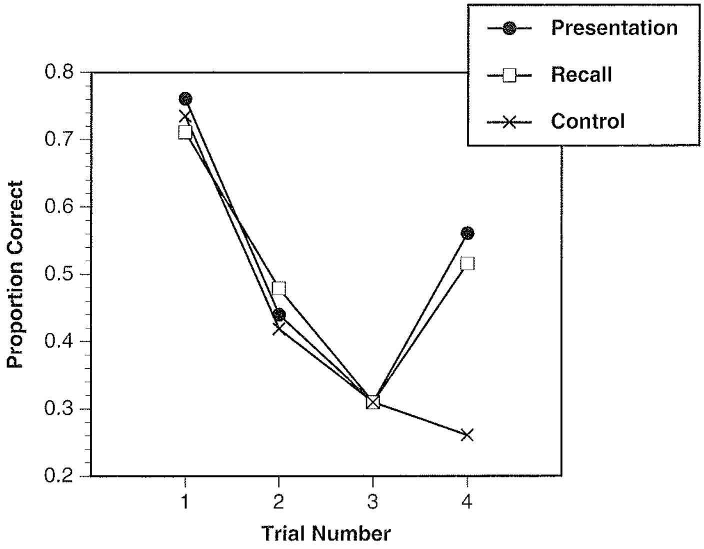

Figure 6.6 Proportion of items recalled correctly on the first four trials of a Brown-Peterson procedure as a function when the cue was presented, either before presentation (the Presentation condition) or after presentation but before recall (the Recall condition). The cue was not given to the control group. SOURCE: From "Retrieval Cues and Release from Proactive Inhibition," by J. M. Gardiner, F. I. M. Craik, and J. Birtwistle, 1972, *Journal of Verbal Learning and Verbal Behavior, 11*, 778–783. Copyright © 1972 Academic Press, Inc. Reprinted with permission.

show release from PI. A second problem is that other manipulations designed to affect the amount of attention do not yield any release (MacLeod, 1975).

A more promising idea is that it is a retrieval effect. Gardiner, Craik, and Birtwistle (1972) examined this hypothesis using two categories of items, each of which could be divided into subgroups. Some subjects received types of games, others types of flowers. Unnoticed by the subjects, the first three trials would all be indoor games (or wildflowers), and the fourth trial would be outdoor games (or garden flowers). All subjects received this subtle change in categories, but only two of the groups were informed of the change. One group was informed prior to presentation of the fourth trial; the other group was informed after presentation of the fourth trial but prior to recall. Nothing extra was said to the control subjects. The results are shown in Figure 6.6.

The control subjects were affected by the buildup of PI, and this continued through Trial 4. If the change is unnoticed, then there will be no release from PI; but both the presentation and recall groups showed an equivalent release from PI. Gardiner and associates argued that this release must be a retrieval effect because that was the only point at which both groups had the category change information at the same time: The recall group could still take advantage of the change, even though they were unaware of this subtle change in category at encoding.

Decay explanations of the results from the Brown-Peterson paradigm do not work. A better explanation is one that views interference effects as arising from distinctiveness.

Relative Distinctiveness

Explanations of forgetting based on *relative distinctiveness* have, as their foundation, the idea that items will be well recalled to the extent that they stand out or differ in some fashion from surrounding items. A well-known example of distinctiveness is the so-called *von Restorff effect.* It turns out that what most people describe as the prototypical experiment and result of this effect differs substantially from what Hedwig von Restorff actually reported in 1933 (see Hunt, 1995). Furthermore, the results were anticipated by Calkins (1896);Wallace (1965) called it the "pre-von Restorff von Restorff effect."

What is most often meant by the von Restorff or isolation effect is the finding that if one item in a list differs from the others, it is better recalled. For example, if the item at position 6 of a 12-item list is red and all the other items are black, item 6—the red item—will be very well recalled. If, on the other hand, all the items are red, then item 6 will not be very well recalled, even though it is the same item in both cases. Pillsbury and Raush (1943) presented three-digit numbers embedded in a list of three-letter nonsense syllables. When there was only one three-digit number, recall of that number was excellent; as the number of three-digit numbers increased and the number of three-letter nonsense syllables decreased, recall of the digits decreased and recall of the nonsense syllables increased.

The most obvious explanation, that the isolated item receives more rehearsal and as a consequence is recalled better, is not borne out empirically (e.g., Einstein, Pellegrino, Mondani, & Battig, 1974). Indeed, given that the effect is observable when the isolate is in the first serial position, explanations based on increased attention to the isolated item are also ruled out.

CogLab **Experiment**

http://coglab.wadsworth.com/experiments/VonRestorff/

The most likely explanation of the von Restorff effect is distinctiveness. The isolated item is distinct, or stands out, from the list of items at the time of recall. This account is similar to the one used to explain the Gardiner, Craik, and Birtwistle (1972) release from PI experiment. In the typical von Restorff experiment, the isolated item will differ in distinctiveness along a size dimension or a color dimension.

Just as items can differ in terms of their distinctiveness along physical dimensions, they can also differ along other dimensions, such as time. This type of retrieval hypothesis suggests an explanation for the data reported by Turvey, Brick, and Osborn (1970). They used a Brown-Peterson task in which consonant trigrams were presented followed by various amounts of distracting activity. Although they had five conditions, we will restrict our attention to only three of them. These three conditions differed only in the duration of the retention interval (how long the subjects performed the distracting activity). One group of subjects counted backward for 10 s, one for 15 s, and one for 20 s. The proportion of items recalled correctly in each condition was equivalent on the first trial (0.85, 0.93, and 0.93, respectively). The proportion of items correctly recalled in each condition was also equivalent on the fourth trial (0.33, 0.30, and 0.30, respectively). Note that these results are already inconsistent with a decay explanation, because more decay should have occurred with the longer intervals.

On the fifth trial, Turvey and associates had every group perform the distracting activity for 15 s; thus, one group's retention interval increased, one group's remained the

Table 6.4 Three groups that receive presentation of the first item (P1) at the same time and presentation of the second item (P2) at the same time; but recall of the second item (R2) occurs after varying delays

	Time								
Group	0	10	20	30	40	50	*P1–R2*	*P2–R2*	*Ratio*
A	P1		P2	R2			30 s	10 s	3:1
B	P1		P2		R2		40 s	20 s	2:1
C	P1		P2			R2	50 s	30 s	1.67:1

NOTE: "Ratio" is the ratio of the time P1–R2 to the time P2–R2.

same, and the third group's retention interval decreased. On this fifth trial, performance was not equivalent: The 10 s group got worse (down to 0.20), the 15 s group remained constant (0.28), and the 20 s group got better (up to 0.38). A change in the duration of the distracting activity is not sufficient to obtain a release from PI, because one group actually performed worse after the change. However, Baddeley (1976) offered an explanation that attributes this result to the relative, rather than the absolute, duration of the distractor activity.

Imagine three groups of subjects who receive two trials in a Brown-Peterson procedure. The first item is presented at Time 0, and the second item is presented after 20 s. Up until recall of the second item, everything is the same in the three groups. The second item is recalled either 10, 20, or 30 s after it was presented, as illustrated in Table 6.4.

Baddeley (1976) calculated ratios for the times involved. The time between presentation of the first item (P1) and the test for the second item (R2) is 30 s for Group A, 40 s for Group B, and 50 s for Group C. The time between presentation of the second item (P2) and the test for the second item (R2) is 10 s for Group A, 20 s for Group B, and 30 s for Group C. These figures and the resultant ratios are shown in Table 6.4. The prediction is that the group with the largest ratio will show the best performance because, Baddeley suggested, at the time of the recall attempt for item 2, recall will be primarily determined by the subject's ability to distinguish the second item from the first item. The closer in time the recall test is to the presentation of the second item and the farther away the recall test is from the presentation of the first item, the better will be the recall of the second item. It is this idea that the ratio is capturing.

The same reasoning can be applied to the Turvey, Brick, and Osborn (1970) data, shown in Table 6.5. The fourth item is presented at Time 0 (an arbitrary starting point).

Table 6.5 Three conditions from Turvey, Brick, & Osborn (1970), represented in the same way as in Table 6.4

	Time											
Group	0	5	10	15	20	25	30	35	40	*P4–R5*	*P5–R5*	*Ratio*
10	P4		P5			R5				10	15	0.67
15	P4			P5			R5			15	15	1.00
20	P4				P5			R5		20	15	1.33

The fifth item is presented either 10, 15, or 20 s later, according to the group. The recall test for the fifth item occurs 15 s later. The same ratios as before are calculated, and again, the larger the ratio, the better the recall of the fifth item.

This example contains a few simplifications. The first three items are not factored in, and the time intervals are not exactly those that appeared in the experiment (we have omitted the time needed to recall). However, this analysis applies even if we factor all these things in; it is just more complex. Similar analyses based on temporal discrimination have been used to explain serial position effects observed in recognition (Neath, 1993b) and in memory for items within a sentence (Neath & Knoedler, 1994).

One particularly interesting phenomenon that this type of view emphasizes is that memory performance can get *better* as time passes (Bjork, 2001). In one experiment, Neath and Knoedler (1994) presented subjects with a series of digitized photographs of snowflakes. These stimuli were chosen because people are unable to assign verbal labels to them, and thus the stimuli cannot be rehearsed verbally (Goldstein & Chance, 1970). These flakes were presented for 1 s on a computer screen with a 1 s gap—or interstimulus interval (ISI)—between one flake and the next. The recognition test occurred after a retention interval (RI) of either 0, 2, or 5 s. For the test, the snowflake shown was either one from the immediately preceding list or one that the subjects had not seen before. Only one recognition judgment was performed after each list. The results are shown in Figure 6.7. With a short retention interval, memory for the first item is at chance levels,

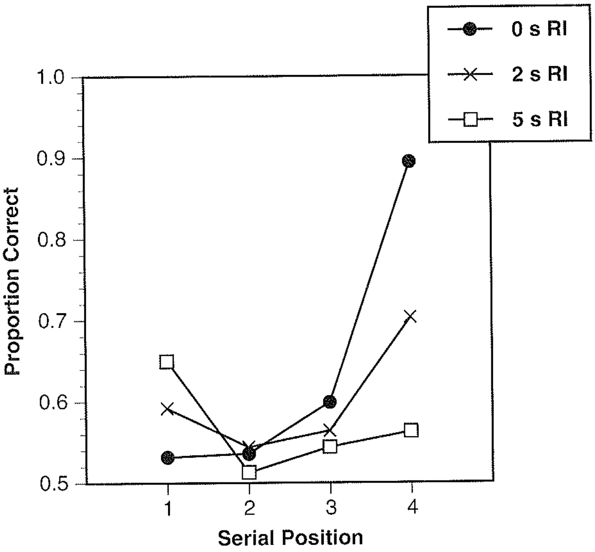

Figure 6.7 The proportion of items recognized correctly as a function of serial position and retention interval. Contrary to predictions based on decay or consolidation theories, there is an absolute increase in performance at the first serial position as the interval between presentation and test increases. SOURCE: From "Distinctiveness and Serial Position Effects in Recognition and Sentence Processing," by I. Neath and A. J. Knoedler, 1994, *Journal of Memory and Language*, *33*, 776–795. Copyright © 1994 Academic Press, Inc. Reprinted with permission.

whereas the last item is almost perfectly recognized. As the retention interval increases, performance on the last item decreases and performance on the first item increases.

A simple model based on distinctiveness explains this result. Murdock (1960) proposed the original idea, which was subsequently refined to include ideas of Baddeley (1976), Bjork and Whitten (1974), and Neath (1993b). The model defines distinctiveness as the extent to which a stimulus stands out from other stimuli, and emphasizes the relation between a given stimulus and the other members of the ensemble. The concept of distinctiveness is thus not applicable to an item in isolation, and the distinctiveness of a particular item can change depending on the other items in the list. As noted previously, for example, if the item at position 6 of a 12-item list is red and all the other items are black, item 6—the red item—will be very distinctive and very well recalled. If, on the other hand, all the items are red, then item 6 will not be very distinctive, even though it is the same item in both cases (von Restorff, 1933; Wallace, 1965).

Distinctiveness can, in principle, be calculated for any set of stimuli that vary along any dimension, whether temporal, physical, or some other dimension, although most of the work heretofore has focused on temporal (time-based) distinctiveness. In the typical memory experiment, the stimuli are equated for concreteness, length, frequency, and so on, and vary systematically only along a temporal dimension. The position of the item along the relevant dimension—here, temporal—is determined by calculating the sum of all the interstimulus intervals (ISI, in seconds) following the item in question, plus the retention interval (RI) following the final item. For example, if the RI is 1 s and each ISI is 1 s, the serial/temporal values (*s*) of a four-item list, beginning with the first item, would be 4, 3, 2, and 1. The final item in the list is a distance of 1 from the recaller, whereas the first item is a distance of 4 from the recaller.

These temporal values then undergo a log transformation, a process based on Weber's Law (see Chapter 1; see also Helson, 1964), yielding the preliminary discriminability of that item (d). For the example given above, the natural log values would be 1.386, 1.099, 0.693, and 0.000, respectively. A distinctiveness value of each item (δ) is calculated according to Equation 6.1, where n is the number of items in the list and k is the serial position of the current item.

$$\delta_k = \sum_{j=1}^{n} \left| d_k - d_j \right| \tag{6.1}$$

All Equation 6.1 really does is calculate a measure of how different each item is from all the other items. We take the absolute value of the difference because we do not care which number is larger, just what the size of the difference is. Thus, δ_1 is equal to (1.386 – 1.386) + (1.386 – 1.099) + (1.398 – 0.693) + (1.386 – 0.000), or 2.367. (Although all values are rounded, the values given as final results are accurate.) We then perform the same calculation for the other items and get values of 1.792, 1.792, and 3.178. (Actually, Neath, 1993b, used a slightly different equation for the middle items, but it is omitted here for the sake of simplicity.) Finally, these values are normalized. This means that they are converted to a scale that reflects the relative size of the numbers. To do this, we add up all the values, obtain a total of 9.129, and then divide each δ value by 9.129 to get the distinctiveness of each item relative to the other items on the list. The resulting values are 0.259, 0.196, 0.196, and 0.348.

What happens to the δ values when the RI is increased? Using the same example, but changing the RI to 5, we get the following serial/temporal values: 8, 7, 6, and 5. The logs of these values are 2.079, 1.946, 1.792, and 1.609. In this case, δ_1 is equal to (2.079 – 2.079) + (2.079 – 1.946) + (2.079 – 1.792) + (2.079 – 1.609), or 0.891. We then do the calculations for the other items and get values of 0.624, 0.624, and 0.989. The sum of these δ values is 3.128, and the normalized values are 0.285, 0.200, 0.200, and 0.316. Notice that with an increase in the duration of the retention interval, the model predicts that primacy will increase (0.259 becomes 0.285) and recency will decrease (0.348 becomes 0.316).

Wright and his colleagues (Wright, Santiago, Sands, Kendrick, & Cook, 1985) have shown increases in performance similar to those found by Neath and Knoedler (1994), but Wright showed that this finding holds for pigeons and monkeys. Knoedler, Hellwig, and Neath (1999) and Surprenant (2001) found that the result also holds for auditory stimuli: As the retention interval increases, memory for the items at the end of the list gets worse and memory for the items at the beginning of the list gets better. Wheeler (1995) has demonstrated, in a different paradigm, increased performance over time when there is no opportunity for further rehearsal or learning, and Bjork (2001) provides a review and extended discussion. Any theory of forgetting must explain not only why memory is often worse as time passes, but also why memory is sometimes better.

Any Evidence for Decay?

The strongest evidence against decay is finding improvement over time. Given the results just described, one might well ask whether there remains any evidence in support of decay. A strong case for decay can be made whenever absolute time has effects beyond those attributable to other sources of forgetting, such as interference.

Baddeley and Scott (1971) tested 424 undergraduates on their memory for a single spoken sequence of three, five, or seven digits after a delay of between 0 and 36 seconds. During the delay, subjects copied letters presented at a rate of 1 per second. The data are shown in the left panel of Figure 6.8.

Baddeley and Scott (1971) argued that the figure shows clear evidence of decay. Proactive interference could not have caused the forgetting, because there were no previous trials. Similarly, retroactive interference from the distractor task could not have caused the forgetting because letters and digits are sufficiently different (Wickens, Born, & Allen, 1963). The final source of interference would be intrasequence effects (Melton, 1963). Baddeley and Scott argued against this view because of the prediction that there should be faster forgetting for longer lists.

It turns out that it is easy to account for the difference in performance between the lists solely as a function of list length. According to the distinctiveness account of forgetting, the more items in the list, the less distinct any one item will be. Furthermore, performance decreases over time not because of decay, but because the memory representations become more similar, and so less distinctive. An extension of the distinctiveness model presented above (Brown, Neath, & Chater, 2002) was fitted to the Baddeley and Scott (1971) data. The right panel of Figure 6.8 shows the same data points, but the lines are those that the model predicts. Thus, these data can easily be accounted for by a model in which there is no decay.

The second study often cited as evidence for decay is one by Cowan, Saults, and Nugent (1997). On each trial, they presented two tones to subjects and asked them to judge whether the second tone was higher or lower in pitch than the first tone. The time between

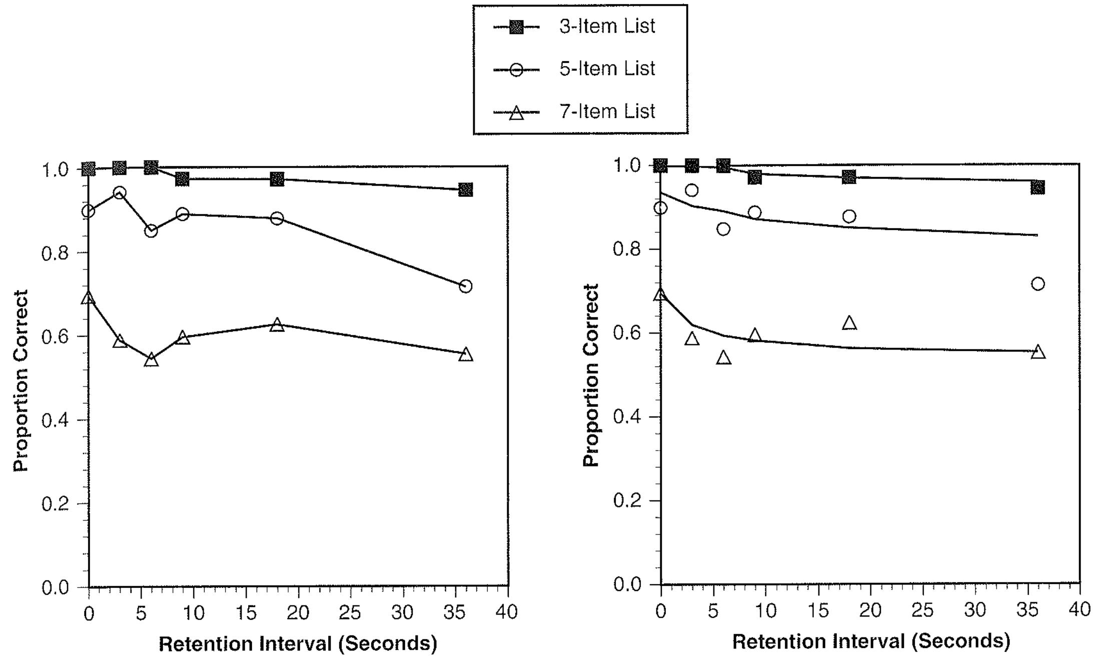

Figure 6.8 Proportion of items recalled correctly on the first trial of a Brown-Peterson procedure as a function of the number of items in the sequence and the duration of the distractor task. When the data points are simply connected, as is done in the left panel, it looks as if there is evidence for decay. The right panel shows the same data points, but this time the lines represent predictions from a distinctiveness model that attributes the forgetting to intrasequence interference: the more items in the list, the more intrasequence interference. SOURCE: Baddeley & Scott (1971).

the tones (the interitem presentation interval, or IPI) varied from 1.5 s to 12 s, and the time between trials (the intertrial interval, or ITI) also varied from 1.5 s to 12 s. With only two items, intraitem interference was minimal. They plotted their data as a function of the ratio ITI:IPI and found that for a given ratio, performance got worse over time. This suggests an effect of absolute time, and so provides evidence for decay. However, Cowan, Saults, and Nugent (2001) later reanalyzed the data, taking into account not only the role of the ITI and the IPI, but also a possible effect of tones on previous trials. They classified some of the trials from the original experiment into those in which the distinctiveness account would predict near optimal conditions, and those with suboptimal conditions. Analysis of the optimal trials found no effect of absolute time on performance. As a result, Cowan et al. (2001, p. 326) concluded that they had "failed to find clear evidence of decay in a situation that has often been viewed as one of the simplest paradigm cases for decay."

Forgetting as Discrimination

One theory of forgetting that can explain better performance as time elapses views memory and forgetting as a discrimination problem. According to Capaldi and Neath (1995), during learning the organism not only processes the material to be learned but also associates that material with a variety of internal and external contextual cues. What ends up being processed is a multidimensional complex of stimuli. Forgetting will occur if the stimulating conditions at test do not sufficiently discriminate between the desired memory and other competing memories, or if the stimuli at test elicit no memories at all. To the extent that the test cues do discriminate, the organism will remember. Thus, forgetting is simply a performance deficit resulting from inadequate stimulus conditions. This approach suggests that forgetting, in the sense of permanent loss, does not occur; there is only a failure to perform because of a difference in the stimulus conditions prevailing at encoding and at test.

The question of permanent loss versus temporary lapse is an important one, but it is difficult if not impossible to answer. If an organism learns something at Time 1, demonstrates memory for it at Time 2, but demonstrates no memory for it at Time 3, has the information been permanently lost, or is this simply a temporary retrieval deficit? In 1929, H. C. Blodgett reported an experiment demonstrating that a lack of performance was at least sometimes not a lack of learning or memory. He placed hungry rats in a complex maze and gave them food when they found their way through. These rats quickly learned the shortest route. He placed a second group of hungry rats in the maze but did not reward them. Not surprisingly, the performance of these animals did not improve over trials. With the introduction of food rewards, however, the error scores of the second group of animals plummeted to that of the first group that had been rewarded with food from the outset. This seminal study illustrates the important proposition that a lack of performance does not necessarily indicate a lack of learning—or, in memory terms, a lack of performance does not necessarily indicate a loss of previously learned information. Whenever an item is not recalled, it is possible that it is only temporarily inaccessible and that with a change in conditions or testing methods, the subject will produce it (Tulving & Pearlstone, 1966).

Loftus and Loftus (1980) took the position that there is permanent loss of information. Their evidence was mostly negative in nature, demonstrating that neither hypnosis (see Chapter 12) nor neurological studies by Wilder Penfield (see Chapter 15) supported the idea of permanent storage. The one positive line of evidence used to support their position came from the misleading eyewitness information paradigm, which we will discuss in Chapter 12. Capaldi and Neath (1995) took the opposite position. Although agreeing with Loftus and Loftus in their analysis of the hypnosis and Penfield studies, Capaldi and Neath suggested that the temporary lapse view made more predictions supported by empirical evidence than did the permanent loss view. In particular, they cited (literally) dozens of studies that demonstrated memory for information previously thought to be permanently lost (see also Miller & Grahame, 1991). To be clear, the temporary lapse view says that if a person (or animal) demonstrates memory for something at Time *N*, then that particular memory will be potentially available at all subsequent times as long as the retrieval cue is appropriate. The lapse view does not say that people remember everything they see or hear.

Two examples will serve here: one in which memory performance became better with the addition of a useful retrieval cue, and one in which memory for various items seems to have lasted at least 50 years. Eich and Birnbaum (1982) asked subjects to learn a list of words in either a sober or an intoxicated state. When tested after a change in drug state, retrieval was poor, as is usually the case (see Chapter 5). However, some of the subjects in the changed state condition were given an additional retrieval cue (such as "Think of all

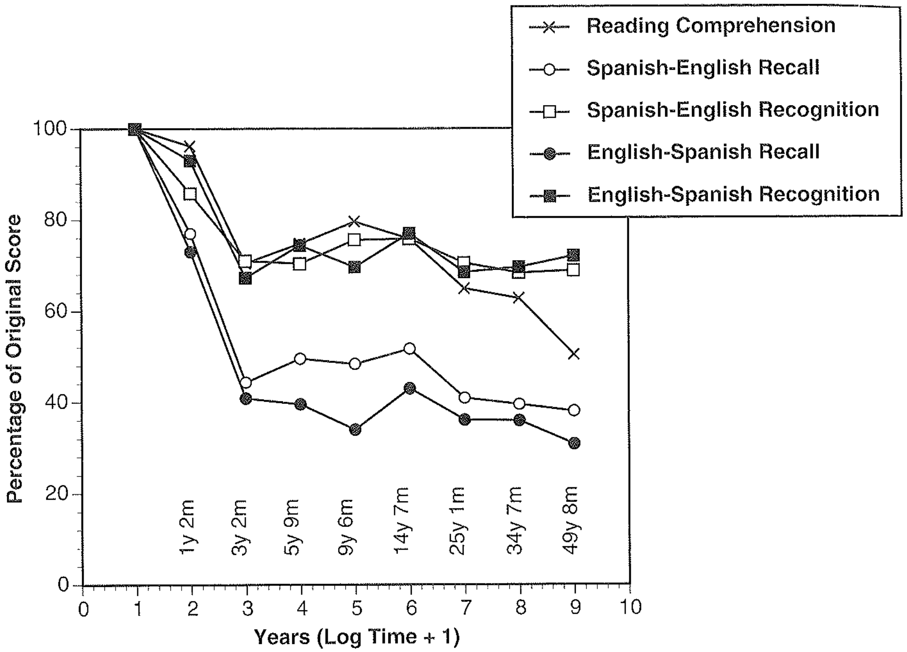

Figure 6.9 Amount of Spanish remembered over intervals from 1 year 2 months (1y 2m) to 49 years 8 months (49y 8m), expressed as a percent of original learning. Any value over 0 indicates that some material was remembered. SOURCE: Adapted from "Semantic Memory Content in Permastore: Fifty Years of Memory for Spanish Learned in School," by H. P. Bahrick, 1984, *Journal of Experimental Psychology: General, 113*, 1–29. Copyright © 1984 American Psychological Association. Adapted with permission.

the flowers that you can"). These subjects were then able to recall many of the items that they previously could not. The temporary lapse view predicts that when the appropriate retrieval cue is presented, evidence of memory—as opposed to memory loss—will be observed. The trick, of course, is finding the right retrieval cue.

The temporary lapse view also predicts that, under appropriate testing conditions, there should be evidence of retention regardless of the delay between study and test. Perhaps the most impressive data on permanent retention come from the studies reported by Harry Bahrick and his colleagues.

Bahrick, Bahrick, and Wittlinger (1975) demonstrated that although memory for people who attended the same high school more than 50 years ago may be very poor when tested with recall, when tested with a cued-recognition test, performance is surprisingly accurate. For recognition of pictures of high school classmates, accuracy remained as high as 80–90% for more than 35 years. Similar feats of retention (in what has been dubbed "the permastore") have been demonstrated in mathematics (Bahrick & Hall, 1991), for city streets and locations (Bahrick, 1983), and for the Spanish language (Bahrick, 1984a). In this last study, Bahrick tested 773 people who had taken Spanish in high school. He collected objective evidence of their original proficiency, including measures of grades, amount of practice, and years of study. He then administered a test to measure current understanding; Figure 6.9 shows the results as a percent of original performance. Although there was a substantial drop during the first three years, there was essentially no further loss over the next 50 years. Ebbinghaus (1885) reported similar results: When relearning certain stanzas from Byron's *Don Juan* after a 30-year interval, he still showed savings.

Each view, permanent loss and temporary lapse, has its proponents and detractors. For the time being, it is worthwhile to remember the basic predictions each view makes. The permanent loss view suggests the information is gone forever, whereas the temporary lapse view suggests that under different conditions with different ways of testing, the information might still be revealed. The results of the studies cited above (see also Capaldi & Neath, 1995; Miller & Grahame, 1991) are predicted by the temporary lapse view but not by the permanent loss view.

Although covered in more detail in the next chapter, results from experiments studying implicit memory also seem to support the temporary lapse view. For example, there is a paradigm known as *repetition priming*. In these experiments, subjects are typically presented with a word fragment that can be completed with only one valid word. For example, *toboggan* might be the target word, and *_o_ogg_n* might be the word fragment; some subjects have recently seen the word *toboggan* on a prior list, and others have not. Repetition priming refers to the finding that those subjects who have seen *toboggan* recently are more likely to complete the word fragment correctly than are those subjects who have not, even if the former subjects are unaware of their prior exposure to the word. Effects of repetition priming can be detected after a delay as long as 16 months in college students (Sloman, Hayman, Ohta, Law, & Tulving, 1988) and after a delay as long as 12 months in amnesics (Tulving, Hayman, & MacDonald, 1991). Such feats of retention after a single exposure are better explained by the temporary lapse than by the permanent loss view.

Chapter Summary

Consolidation theory attributes forgetting to problems that occur at storage. It predicts that if the memory trace is not consolidated, it will never be recalled. However, many studies show that a change in retrieval conditions is often sufficient to elicit a memory that was supposedly not consolidated. Classic interference theory attributes forgetting to response competition, altered stimulus conditions, and set. Both proactive interference (items you already have in memory can interfere with your ability to remember recently learned information) and retroactive interference (recently learned information can interfere with remembering items you already have in memory) can cause substantial decreases in memory performance. One of the few areas of memory in which decay theory is invoked is short-term/working memory. Decay theory has to predict that memory will always get worse over time. Many findings from the Brown-Peterson and other paradigms disconfirm this prediction. Current theories of forgetting view memory as a discrimination problem in which items will be recalled well if they are distinctive or stand out from competing items at the time of retrieval. This type of view is consistent with the processing accounts discussed in the previous chapter.